R E P O R T R E S U M E S

ED 002 943 24

INTERPERSONAL PERCEPTION--THE EFFECT OF TRAINING IN PERCEPTUAL THEORY, OBSERVATION AND ANALYSIS OF BEHAVIOR UPON ACCURACY OF PREDICTION OF CHILDREN'S SELF-REPORTS.

BY- GORDON, IRA J. SPEARS, WILLIAM D.
FLORIDA UNIV., GAINESVILLE

REPORT NUMBER CRP-813 PUB DATE 31 JAN 62
EDRS PRICE MF-$0.09 HC-$1.24 31P.

DESCRIPTORS- *PERCEPTUAL DEVELOPMENT, PERCEPTION TESTS, COLLEGE INSTRUCTION, *INTERPERSONAL PROBLEMS, *SELF CONCEPT, BEHAVIOR PATTERNS, *STUDENT BEHAVIOR, TRAINING, *PREDICTIVE ABILITY (TESTING), INDIVIDUAL DIFFERENCES, GAINESVILLE, FLORIDA

THE PURPOSE WAS TO SUBSTANTIATE THE HYPOTHESIS THAT TRAINING COLLEGE STUDENTS IN SELF-PERCEPTUAL THEORY, IN OBSERVING CHILD BEHAVIOR, AND IN ANALYZING BEHAVIOR FROM A PERCEPTUAL ORIENTATION INCREASES THEIR ABILITY TO INFER FROM THE BEHAVIOR RECORDS OF A CHILD HIS STATEMENTS ABOUT HIMSELF. THREE GROUPS OF SUBJECTS, ONE EXPERIMENTAL AND TWO CONTROLS, PARTICIPATED, MAKING UP A TOTAL SAMPLE OF ABOUT 230 STUDENTS. APPROXIMATELY 70 OF THESE (ASSIGNED TO THE EXPERIMENTAL GROUP) RECEIVED ADVANCED TRAINING IN A HUMAN DEVELOPMENT COURSE WHICH EMPHASIZED THE COMBINATION OF PERCEPTUAL THEORY AND ANALYSIS. EACH GROUP WAS GIVEN A PRE- AND POST-TEST TO MEASURE ITS ABILITY TO INFER SELF-PERCEPTIONS OF A CHILD IN ACCORDANCE WITH THE PROCEDURE DESCRIBED IN THE PROJECT HYPOTHESIS. CORRELATIONS WERE COMPUTED FOR ALL STUDENT RESPONSES, AND AN ANALYSIS OF VARIANCE DESIGN WAS USED TO TEST THE EFFECT OF TRAINING UPON ABILITY TO INFER. THE STUDY HYPOTHESIS WAS NOT SUBSTANTIATED BY THE RESULTS. THE PROJECT, HOWEVER, DID LEAD TO THE DEVELOPMENT OF TWO INSTRUMENTS WHICH RELIABLY DETECTED INDIVIDUAL DIFFERENCES IN ABILITY TO INFER SELF-REPORTS FROM THE MIMEOGRAPHED OBSERVATIONAL DATA. THUS, FOR THE PURPOSE OF FURTHER RESEARCH, THE NEED FOR A SHIFT IN APPROACH WAS IMPLIED--FROM THE ASPECT OF INTENSIVE GROUP TRAINING TO THE POINT OF CONCENTRATING EFFORT ON THE INDIVIDUAL PERCEIVER OR INFERRER OF BEHAVIOR AND CHILD'S SELF. (JH)

ED002193

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

242

INTERPERSONAL PERCEPTION: THE EFFECT OF TRAINING IN PERCEPTUAL THEORY, OBSERVATION AND ANALYSIS OF BEHAVIOR UPON ACCURACY OF PREDICTION OF CHILDREN'S SELF-REPORTS

Ira J. Gordon, University of Florida

William D. Spears[1], University of Florida

January 31, 1962

---

[1]Now at Auburn University, Auburn, Alabama

3414a

Statement of the Problem

Current educational thought emphasizes the importance of the constructs of perception, self concept and self perception at all levels of educational practice. For example, Havighurst (1959) states that a major cause of underachievement in gifted children is their inability to see themselves as capable. Texts in education such as Kelley, Education for What is Real (1947) and Combs and Snygg, Individual Behavior (1959) emphasize the necessity for teachers and other school personnel to understand the self concepts and perceptions of youngsters in order to develop appropriate curriculum experiences for them. In the field of guidance, Arbuckle (1957), Gordon (1956), and Moustakas (1956) among others suggest the importance of the pupil's self as the major focus for teachers who are attempting to implement guidance and mental health practices in the classroom. To some degree, authorities assume that teachers are capable of making valid and accurate judgments as to the self concepts of their pupils, such judgments being based mainly upon teacher observation of pupils' classroom behavior.

It has not been clearly demonstrated through research that teachers are capable of making valid judgments about the self of children. In particular, there is currently no published research on the effects upon adequacy of judgments of the pre-service training of teachers from the perceptual, self concept point of view. Research, therefore, is needed to study the problem of educating teachers to observe and to infer children's perceptions and conceptions of self with accuracy.

The present study offers one empirical approach to this problem. It was designed to test the feasibility of a perceptual approach in enabling teachers to make valid inferences of children's self estimates. If this and supporting studies demonstrate that this is feasible, then undergraduate courses in human development can have as one of their goals the achievement of this ability by students.

The problem investigated in this study then was whether or not a course emphasizing a particular theoretical viewpoint, together with practice in observation and analysis based upon that viewpoint, increases the ability of undergraduate teacher education students to infer accurately the self-reports of children.

Review of the Literature

In order to investigate this problem several factors needed to be considered. First, what technique or instrument is most appropriate in securing information as to the self-concept of children? Second, what variables might influence the student's accuracy of interpersonal prediction? The latter problem includes two facets, the educational experience and the personality of the education student.

Measure of Self-Concept

The use of self-reporting devices is a common procedure for gathering data on children. Pauline Sears (1960) and Perkins (1958) both use "self-report" and "self-concept" synonymously in their studies of preadolescent children. Perkins used a Q-sort technique based upon the Jersild (1952) categories. Sears developed a 100 item, 5-point rating scale. Both the Sears and Perkins studies indicate that the level of accuracy of teachers in estimating children's self-evaluations is not very high. The highest correlation in Sears' study, for example, was between teacher judgment and boys' self-concepts of their mental ability, the correlation being of the order of .50.

A review of the literature by Gordon and Combs (1958) revealed that the major techniques of other investigators of the self-concept were either Q-sort or self-rating scale devices. Combs and Soper (1957), however, carefully point out the problem of confusing a child's self concept with his self report. They indicate that such a variety of social variables and perceptions influence the self report that the report may not really be a child's self-concept.

There is little room for disagreement with Combs and Soper. It is common knowledge that much of what we say about ourselves is either untrue or distorted. Furthermore, self-concept has evolved into a formal construct which purports to explain the <u>why</u> of behavior, not just denote it. It must be recognized, however, that subjects are frequently capable of giving accurate reports of their feelings and that a carefully designed self-rating scale can tap these areas of self knowledge.

Lipsitt (1958), Coopersmith (1959), and Brandt (1958) have shown that self-reports are related to other measures of the self-concept. In addition, Allport (1953) urges that researchers assume that the healthy person's direct statements about himself can be virtually taken at face value.

The instrument used in this study to measure self-concept, therefore, is a self-rating scale based upon the Jersild (1952) categories.

The problem of the accuracy of interpersonal prediction in the area of person perception has been investigated by Cronbach (1955) and Crow (1957) as well as others (Tagiuri and Petrullo, 1958). For the purposes of this study, the conclusion of Cronbach and Crow, that training programs which merely provide for exposure and observation might tend to decrease the accuracy of the trainee's responsiveness to individual differences because the trainee becomes more aware of individuality, was a key notion. In their separate studies they point out that if the subject increases his differentiation without a corresponding increase in ability, he will make a greater number of errors.

Courses in human development characteristically utilize normative approaches in which students can come to see the behavior of a particular child in relation to the behavior of age mates. The perceptual orientation focuses, however, upon individuality. It might very well be that such focus without corresponding training in utilization of this position would have the effect suggested by Cronbach and Crow. The training program, therefore, was designed to attempt to overcome the difficulties suggested by these researchers. Such an approach would resemble the in-service child study program advocated by Prescott (1957), participation in which Perkins (1958) found was a significant variable in improving the accuracy of teacher judgment.

When we turn to the question of the personality structure of the judge or, in this study, the education student, Baker and Block's (1957) study of accuracy dealt mainly with the problem of the ego control of the judge rather than with the nature of his training. It has long been acknowledged that the judge himself is an important variable in the judgment. This present study did not attempt to investigate this variable. The personalities of the individual education student were not examined, but the attempt was made through statistical means to equate the judgment of experimental and control groups before training.

Hypotheses

The central hypothesis of this study was that training students in perceptual self theory, in observing child and adolescent behavior, and in the analysis of such behavior from a perceptual orientation, would increase the ability of students to infer accurately a child's statement about himself. Specifically,

a. Students receiving training in the combination of theory and analysis will increase significantly in accuracy from before to after training.

b. The increased accuracy of students receiving training in a combination of theory and analysis will differ significantly from the increase, if any, of those with no training in either.

c. The increased accuracy of students receiving training in the combination of theory and analysis will differ significantly from the increase of those receiving training only in observation.

Procedure

Sample

The subjects for this experiment were selected as follows: the experimental group included 69 students who registered for a total of three sections of an upper division (open to juniors and seniors) course in Human Development during the spring semester, 1960, and the fall semester of 1960-61. Students could elect this course after completing a prerequisite course in Human Development at the University which includes training in observation but does not necessarily include perceptual theory or any analysis of behavior.

Students who transfer to the University and who have had a previous course in child or adolescent psychology or a course in educational psychology or human development, but lack training or experience in observation, are required to take this course. The experimental classes were thus composed of students who had a previous course in the content of human development but were mixed in terms of previous observational experience.

The control groups were selected so as to include a number of subjects who had taken previous course work, including observation experience with children, and subjects who did not have such experience, so that they would be comparable in background to the experimental group subjects. In addition, control subjects were selected so as to represent different levels of training, i.e., sophomore, junior, and senior, and different areas of specialization, i.e., elementary, secondary students.

The upper division (junior and senior) control subjects were taken from four classes offered during the spring semester of 1960. Two of these classes were sections of a required general methods course for elementary majors and two were a required general methods course for secondary majors. All students who registered for these four sections were used. This upper division control group consisted of 57 subjects after those who were also enrolled in the experimental classes were eliminated.

In addition, a total of 101 control subjects were taken from three sections of an introductory (sophomore) class in social foundations of education and from one section of the introductory course in human development. This group included both those with previous observation experience and those who were just beginning a course in human development and observation experience. Measures on all subjects were obtained during the spring semester 1960 and the fall semester 1960-61.

The sample thus consisted of people with and without observation experience, with and without content experience in human development, at varying levels of training, and with differing occupational aspirations within the field of teaching. The interactive effects of the experimental course therefore with these variables could be determined.

The Training Program

The experimental group was given a three semester hour course in Dynamics of Human Development. The essential features of this course, from the point of view of the experiment, was the laboratory in observation and analysis of behavior, and the content which was centered around the perceptual and self-concept theoretical position. A description of the content and procedures follow.

Observation. Each student was assigned to an observer team consisting of two to four students, with most groups having three members. These teams were then assigned to study, through observation in the school setting only, the behavior of a particular pupil in the P. K. Yonge Laboratory School. Students who planned to become high school teachers studied a high school pupil; those planning for elementary education observed in the elementary division of the Laboratory School. Wherever possible, students were assigned to the grade level of their choice.

The pupils selected for observation were chosen by their teachers. Each teacher submitted the names of five "average, typical" children at the beginning of each school year and these were the only children studied by the various groups in Education and Medicine who used the Laboratory School.

During the semester each member of the observer team made ten one-hour observations on the same pupil, on other than class time. He kept a specimen record, taking notes throughout the observation. This meant each three-person team had thirty hours of behavioral data to use for analysis and inference as to the self-concepts of the pupil being observed.

In addition, they had access to the health records, could arrange interviews with teachers and, in several instances, were able to gather sociometric information. They were not able to visit the home, interview the pupil, or talk with the parents. Standardized test data were often available to them during the teacher interview.

Analysis. The observer teams met regularly, sometimes during class, more often on their own time, to discuss their observations, categorize their data, and make inferences as to the perceptions of the pupil being observed.

The first analytic step was simply a check of observer objectivity. One observation of the child was made by all members of the team simultaneously, and they checked each other for reliability, amount of detail, and other criteria for good observation.

The second step consisted of two types of inference-making. First, the categorization of the data into the general classes of school, peer, physical and self. Each event was coded as one of these four depending upon whether the individual observer, and then his group, saw the data as

providing information about this aspect of the environment of the pupil. For example, descriptions of teacher behavior were coded school; peer behavior, other than that of the observee, were coded peer; such data as physical appearance, height, weight (but not grooming) were coded physical; and all the behavior of the observee were coded self regardless of the target or the setting. These large areas were subdivided (see Appendix A) so that the team could develop a clear picture of both the psychological and sociological setting in which the observee's behavior was occurring, and a clear idea of his behavior in each of these various settings. To some degree, this procedure represents an adaptation of both Prescott's (1957) child study approach and Barker's (1954) psychological ecology method.

The following is an excerpt from a record on a twelfth grade pupil showing the coding system. For the purposes of this illustration, only the major codes (I for school, II for peers, IV for physical, and V for self) are indicated.

I. "Mr. Brown put a sample pulley problem on the board. Jane watched the board the whole time Mr. Brown worked the problem. Finally when he was almost through, he asked Jane, 'How does this problem work out to get a one pound force?'

V. "Jane said, 'I can't explain it but I can do it!'

II. "One of the girls sitting behind her said, 'Oh, go ahead!'

I. "The bell rang and Mr. Brown said, 'Saved by the bell, no not quite.'

V. "Jane repeated, 'I can do it but I can't explain it.'

I, II. "Mr. Brown called on another girl and she said, 'Three pounds, anyone can see that.'"

The second type of inference was the one that would be most related to the outcome hypothesized in this study. For each event, the individual observer and then his team made what were labelled "perceptual inferences," that is, once behavior had been coded as "self," and with the basic assumption that behavior is a function of perception, the student inferred how the event was perceived by the behaver, and how the behaver perceived himself. This was an extremely difficult task for these students, due partly to inexperience and possibly to the emphasis upon objectivity in other courses they had taken. They also realized how much their own perceptions might be coloring not only their inferences but also their supposedly objective observations. Gradually, the team members found that one of the strengths of the team was the balance provided by the varying perceptions as to the meaning of behavior. For example, the observer made the following inferences alongside the above observation excerpt:

1. "Why did he have to ask me?

2. "I can see how to do it but I'm not sure enough of myself to attempt it.

3. "I wish I were saved by the bell.

4. "I knew it I guess. Oh well, so I didn't participate again."

It will be noted that these are not "depth" inferences but surface statements. As the students progressed, they were able to make more substantial inferences.

Before observations began, mimeographed case materials which were not to be involved in testing were used for training in coding and inference making. Several class laboratory exercises were conducted to clarify the coding scheme, train for accurate observations, and introduce the students to the inference process.

In addition, the team was asked to develop a plan, after their first four observations, as to the ways in which they would use their remaining observations, teacher interviews, and health data to test their inferences. This constituted a section of the preliminary report (see Appendix B) submitted at that time.

The students' final reports (Appendix C) were devoted mainly to analysis of their data in terms of the self, with the environmental aspect receding.

Subject matter. The first unit was on perceptual theory, the relationships between perception and behavior, the concept of the self. Subsequent topics were developed in each of the three experimental sections on the basis of pupil-teacher planning. They all included, although in differing ways, age-graded data from infancy through adolescence, teacher-pupil relationships, the development of language and concepts, and many standard child and adolescent psychology materials. All were discussed, however, from the viewpoint of their meaning to the child, and of how the teacher could understand and apply these data to his work with pupils.

Role-playing, large and small group discussion, and lectures were all used to convey subject matter.

In addition, each student selected a term paper topic, wrote the usual type individual library research type paper, but then was required to add several pages discussing the personal meaning the paper held for him. In this and all other learning situations, the importance of personal meanings as a motivational force was emphasized. Such emphasis, it was hoped, would add to their sensitivity about the personal meaning of behavior to the pupil being observed.

The Instruments

The behavior record. The students previously enrolled in the upper division human development course in the spring of 1959 observed the behavior of youngsters in the Laboratory School and compiled behavior records in the same fashion that the training group later did in the experimental classes. Teams of three observers studied pupils for 12 hours for each observer for a total of 36 hours for the team. In order to infer the self-concept of these pupils, these students observed the classroom and other aspects of the school situation in order to seek as wide an array of the pupil's behavior as it was possible to do within the school setting.

The behavior records for this study were compiled by the senior investigator from these detailed 36-hour observation records. The records for approximately 10 hours of observation on each of three pupils, Debbie a third grader, Alice a sixth grader, and John a twelfth grader, were selected from the 36 hours in each. The attempt was to include in the behavior record for each of these pupils sufficient variety of behavior to allow the judges to infer the self reports of these three pupils. Ten hours rather than the total record was used because of the limitations of time and also because presenting a judge the whole unedited record would have made his task exceedingly confusing and difficult.

Self-rating scale. A forty-item, 5-point rating scale, "How I See Myself," was developed during the academic year 1958-59. This scale was designed to tap self-estimates of various areas of the self such as the physical body, peer relations, academic achievement, emotional control. The first draft of this scale was revised in the spring of 1959 and administered in all grades from three through twelve at the P. K. Yonge Laboratory School. A test for reliability was run, and reliability was sufficiently high to indicate justification of the use of this instrument for research purposes. The three pupils, Debbie, Alice, and John, along with their classmates completed this scale during the same semester in which the education students were observing and recording their behavior and developing the behavior record.

The Measure of Inference of Self Report

At the beginning of the semester during which the subjects participated in the study, each subject in the experimental and control groups completed the "How I See Myself" questionnaire as they thought each of the three pupils would complete it. Their only source of information about the pupils were the three observation records described above. The subjects completed this questionnaire for each pupil a second time at the end of the semester during which they participated in the study. It took the subjects approximately 2½ hours to read the three records and complete the three rating scales.

The measure of ability to infer self reports was derived by correlating the pattern of ratings assigned by the subjects with the pattern of rating given by the pupil, vis.

$$r_{jk} = \frac{(X_{ij} - \bar{X}_{.j})\ (X_{ik} - \bar{X}_{.k})}{N\ \sigma_j\ \sigma_k}$$

where

r_{jk} = the correlation between the pattern of ratings assigned by the jth subject and the pattern of ratings assigned on the kth pupil;

X_{ij} = rating assigned on the ith item by the jth subject;

$\bar{X}_{.j}$ = mean rating on all 40 items for the jth subject;

X_{ik} = rating assigned on the ith item by the kth pupil;

$\bar{X}_{.k}$ = mean rating on all 40 items for the kth pupil;

N = 40 = number of items on the scale;

σ_j = standard deviation of ratings on all items for the jth subject;

σ_k = standard deviation of ratings on all items for the kth pupil.

Three r's, one with each pupil, were obtained for each subject at the beginning and again at the end of the semester during which they participated in the study. The r's were then transformed into Fisher's z's for use in the analyses below.

Results

Table 1 presents the means and standard deviations of z's on the measures of inference of self-reports for each group on each of the three pupils before and after training. The N's are less than those of the original sample due to dropouts and absences from class at the second testings, and because all subjects who were not planning to become teachers were excluded. Also, the N's for individual groups vary from one measure to another because some subjects completed questionnaires for some pupils but not for others.

No consistent trends before and after training are observable in the means. The mean z's for correlations with pupils Debbie and John tend to be similar, while those with pupil Alice are substantially higher. The standard deviations are smaller at the end of the semester than at the beginning for seven groups and measures. The two exceptions are the standard deviations for upper division controls on pupils Debbie and John.

TABLE 1

Means and Standard Deviations of z's on Three Measures of Inference of Self-Reports for Three Groups Before and After Training

Instrument:	"Debbie" N	Mean z	σ	"Alice" N	Mean z	σ	"John" N	Mean z	σ
Group:									
Experimental	57			52			53		
Before		.128	.227		.260	.293		.146	.154
After		.112	.148		.291	.164		.100	.146
Upper Div. Control	47			42			40		
Before		.129	.159		.321	.208		.110	.186
After		.157	.168		.283	.145		.078	.187
Lower Div. Control	69			62			61		
Before		.198	.164		.316	.204		.117	.195
After		.134	.153		.317	.195		.094	.180
TOTAL	173			156			154		
Before		.156	.189		.299	.240		.125	.180
After		.132	.157		.299	.173		.092	.171

An analysis of variance schema for the overall analysis of data is presented in Table 2. There are three main effects, time, instrument (pupil), & group and one pseudo-effect, individuals within groups. Only 37 subjects in the upper division controls had completed all three measures, so cases were eliminated randomly from the other groups until each group had 37 cases.

TABLE 2

Analysis of Variance Schema for Overall Analysis

		Before Training			After Training		
Instrument:		"Debbie"	"Alice"	"John"	"Debbie"	"Alice"	"John"
	Subject						
Group:							
Experimental	1						
	2						
	3						
	.						
	.						
	.						
	n						
Upper	1						
Division	2						
Control	3						
	.						
	.						
	.						
	n						
Lower	1						
Division	2						
Control	3						
	.						
	.						
	.						
	n						

The analysis of variance data are shown in Table 3. The only significant components were instruments ($F = 52.57$, $p < .001$) and the instrument by subject interaction ($F = 1.45$, $p < .05$). The component of most interest in this study, time by group interaction, was not significant ($F < 1$). This component is a measure of differential changes in groups with time, and thus a direct test of the principal hypothesis that the experimental group would improve more than the controls.

TABLE 3

Analysis of Variance Data Using All Instruments

Source	Σx^2	df	s^2	F
Times	.03875	1	.03875	1.48
Groups	.07979	2	.03990	<1
Instruments	4.45596	2	2.22798	52.57**
Subjects	5.89409	108	.05457	no error term
T X G	.00831	2	.00416	<1
T X I	.03032	2	.01516	<1
T X S	2.83571	108	.02626	<1
G X I	.18313	4	.04578	1.44
I X S	9.15534	216	.04238	1.45*
T X G X I	.12698	4	.03174	1.09
Residual	6.30633	216	.02920	
TOTAL	29.11471	665		

*Significant at or beyond the .05 level.
**Significant at or beyond the .001 level.

No estimate of the reliability of the instruments was obtained. It was decided, therefore, to repeat the analysis with each measure separately in order to test for significance of differences among subjects. The schema is the same as that shown in Table 2 except that the measure classification is missing.

One complication in this procedure for testing for reliability should be pointed out. Although McNemar (1955, p. 290f) derives a test for significance of reliability from a variance design in which subjects and time of testing are the main effects, Spears (1954) has shown that the introduction of a variable (e.g., training) between the first and second testing will lower the consistency of scores obtained from two testings if the introduced variable has different effects upon subjects. Hence, the F's for subjects given in Tables 4a, 4b, and 4c probably underestimate the significances of the reliabilities for single testings.

TABLE 4a

Analysis of Variance Data for "Debbie"

Source	Σx^2	df	s^2	F
Times	.01730	1	.01730	<1
Groups	.18569	2	.09284	2.48
Subjects	4.04196	108	.03743	1.59*
T X G	.07453	2	.03726	1.58
Residual	2.54288	108	.02355	
Total	6.86236	221		

*Significant at .01 level.

TABLE 4b

Analysis of Variance Data for "Alice"

Source	Σx^2	df	s^2	F
Times	.00030	1	.00030	<1
Groups	.04908	2	.02454	<1
Subjects	6.24029	108	.05778	1.41*
T X G	.06002	2	.03001	<1
Residual	4.42678	108	.04099	
Total	10.77647	221		

*Significant at the .05 level.

TABLE 4c

Analysis of Variance Data for "John"

Source	Σx^2	df	s^2	F
Times	.05146	1	.05146	1.04
Groups	.02815	2	.01408	<1
Subjects	1.58906	108	.01471	<1*
T X G	.00074	2	.00037	<1**
Residual	5.35052	108	.04954	
Total	7.01993	221		

*Inverted F = 3.37, significant beyond the .001 level.

**Inverted F = 133.9, significant beyond the .02 level.

The results of the separate analyses are shown in Tables 4a, 4b, and 4c. For Debbie, individuals differ significantly at the .01 level (F = 1.59). No other component is statistically significant.

For Alice, subjects differ significantly beyond the .05 level (F = 1.41). The group and the time by group components are not significant. The time difference is extremely small, but the inverted F is not significant.

A peculiar result occurred for John. The only significant F's obtained were from the inverted ratios for subjects and for the time by group interaction. The inverted F's were 3.37 (df's = 108 and 108) for time and 134.44 (df's = 108 in numerator and 2 in denominator) for the T X G interaction. The first result means that total scores, before and after, on John varied significantly less from subject to subject than separate scores for individual subjects before and after training. The second significant inverted F means that the changes in scores for groups from before to after training were more consistent from group to group than were the changes of subjects within these groups.

The former result could be accounted for if the subjects rated John according to some common "textbook" interpretation of the observation record which was taught to some extent in all classes used in this study. To be valid, however, this hypothesis requires a reduction in variation of measures from before to after training. Furthermore, the greatest reduction in variability should be for the lower division students. On the contrary, the summary data presented in Table 1 reveals that one group, the upper division control, remained constant in variability while the other two decreased only slightly. No further explanation of these two oddities is forthcoming, except that they are possibly rare chance occurrences.

One further analysis of the above type was run for Debbie. Ten more subjects per group could be included since it was not necessary that data on all instruments be available. The results, shown in Table 5, are not appreciably different from those shown in Table 4a. Individuals differed significantly and no other component was significant.

TABLE 5

Analysis of Variance Data for "Debbie" with Ten Additional Subjects per Group

Source	Σx^2	df	s^2	F
Times	.0176	1	.0176	<.1
Groups	.1148	2	.0574	1.69
Subjects	4.6985	138	.0340	1.37*
T X G	.0771	2	.0386	1.56
Residual	3.4286	138	.0248	
Total	8.3366	281		

*Significant at the .05 level.

A second overall analysis was made using the variance design shown above in Table 2. The measures on "John" were excluded from this analysis, however, because of the peculiar results revealed in Table 4c. Otherwise, the second overall analysis was the same as the first.

The results are presented in Table 6. The only significant component was instruments, with an F 56.33 and $p < .001$. As before, this result reveals only that the correlations of patterns of ratings assigned by subjects and by pupils were significantly greater for "Alice." The Instrument by Subject interaction was almost significant, with an F of 1.31 only slightly less than the value of 1.38 required for significance at the .05 level.

No further analyses were made. The failure of the Time by Subject interaction to reach statistical significance ($F < 1$) indicates that, since subjects did not change differentially with time, no a priori classification of subjects would be related to change. Hence, differences between subjects classified according to whether or not they had previously taken courses requiring observations of children, according to areas of major interest, etc., cannot be expected to be statistically significant.

TABLE 6

Analysis of Variance Data for Overall Analysis
"Debbie" and "Alice"

Source	Σx^2	df	s^2	F
Times	.00651	1	.00651	1
Groups	.21084	2	.10542	2.09
Instruments	2.51854	1	2.51854	56.33*
Subjects	5.45323	108	.05049	No error term
T X G	.00879	2	.00440	1
T X I	.01110	1	.01110	1
T X S	3.27210	108	.03030	1
G X I	.02393	2	.01196	1
I X S	4.82903	108	.04471	1.31
T X G X I	.12576	2	.06288	1.84
Residual	3.69754	108	.03424	
Total	20.15737	443		

*Significant beyond the .001 level.

Discussion

The results indicate that our hypotheses were not substantiated. Students in the experimental group did not demonstrate any gain after training. They did not demonstrate any gain over the control groups. There was no relationship between previous observational experience or previous content experience and ability to infer self-reports.

How can we account for this? The first and most obvious explanation is that the theoretical position is not valid. However, no experiment such as this is a direct test of a theory.

The second consideration is in the training itself. It may be that this particular type of training is not able to develop sufficient sensitivity to the inner world of the pupil. Too much stress may have been placed on the objective elements of observation and coding and not enough on overcoming the initial sense of inadequacy about perceptual inference-making. Not enough attention may have been given to subjective experience.

It may be that the training design is adequate, that one three-hour course is insufficient. More time may be needed to develop skill in making inferences. It may also be that the re-test was administered too soon after training, before students had a chance to assimilate and integrate their experiences.

We can look at the training provided in another light. There may be a theoretical problem here of some importance. The basic assumption used was that behavior is a function of perception and perception is a function of self-concept. In order to infer self-concept, then, one studies behavior. Students were trained to infer perception and then self-concept from behavior. Their success in doing this, however, was measured by their ability to infer pupils' self-reports from behavior. The relationship between self-report, self-concept, and behavior at best may be a very tenuous one. Some recent work by Combs, Soper, and Courson (1961) found no significant correlation between pupils' self-reports and observer inferences of self-concept. Their observers had only one and a half hours of observation on pre-adolescents but were able to talk with the subjects.

This lack of correlation may have been due to lack of sufficient behavioral data. For example, both Coopersmith (1959) and Sears (1960) found a significant relationship between teacher estimate and self-estimate. No correlations, however, were very high.

This raises the possibility of the limits of correlation between self as measured by inference from behavior and self as reported. Rather than engage in fruitless debate as to which one is "really" self, it may help to conceptualize a model of the self in which the access route used is seen as revealing only a facet of the total organization. This facet, real in itself, is not the whole self, nor does it overlap with other facets to any great ex-

tent. Different research approaches tap different phases of the self, each growing out of the central core, but each yielding different types of data. Such a model, very crudely, may look something like Figure 1. This will be discussed further in Implications.

Two cases were found to differentiate between subjects. Evidently no results were related to academic training, but did indicate that students differ in their abilities to perceive some measured aspects of the self of another person, and that the two instruments reliably detect this ability.

Why did these cases differentiate and John fail to do so? Alice produced the highest correlations, Debbie was the most reliable instrument. Debbie has the most data (20 pages), Alice has 13 pages, and John 11½. Debbie and Alice are third and sixth graders respectively, John is a twelfth grader. It may be that all three of these variables (amount of data, age, and sex of self-reporter) contribute to the differential effectiveness and reliability of these instruments.

Previous academic experience, and the training itself, did not seem to contribute to the differences between individuals in their ability to infer. Whence, in addition to the variables residing in the instruments, came the difference? Certainly, the observer himself suggests itself as the next logical place to explore. The Baker and Block (1957) study indicates the feasibility of investigating the observer. Such constructs as "observer empathy" and "self-disclosure," for example, are worth investigating.

Implications

Previous studies of pupil-teacher perceptions have been bound to the classroom because of the need to compare teachers with their own pupils. The cases of Debbie and Alice offer a beginning break in this bondage. Students tested on Alice do almost as well as did the classroom teachers on their pupils in the Sears 1960 and Perkins 1958 studies. The correlations on Debbie are lower. Perhaps it now is possible to measure undergraduates' sensitivity to personal meanings of pupils, or compare populations without restriction to one location, permitting study of abstractions with fewer of the extraneous influences characteristic of normal classroom studies and the necessity to use different classes, often in different schools. The use of control groups becomes more feasible. The utilization and further development of these cases is thus the first implication. With the release from the restrictions of location and actual classroom observation, studies comparing the relative abilities of teachers with non-teachers can be made; studies can be conducted to explore whether these instruments are at all predictive of pupil-teacher rapport, using as the measure of rapport either the Davidson and Lang (1960) procedure or the OScAR technique of Medley and Mitzell (1959, 1961). Does possession of the ability to infer self-attitudes relate to teacher effectiveness? The development of "Debbie" and "Alice" offer at least one way to attack this important educational question.

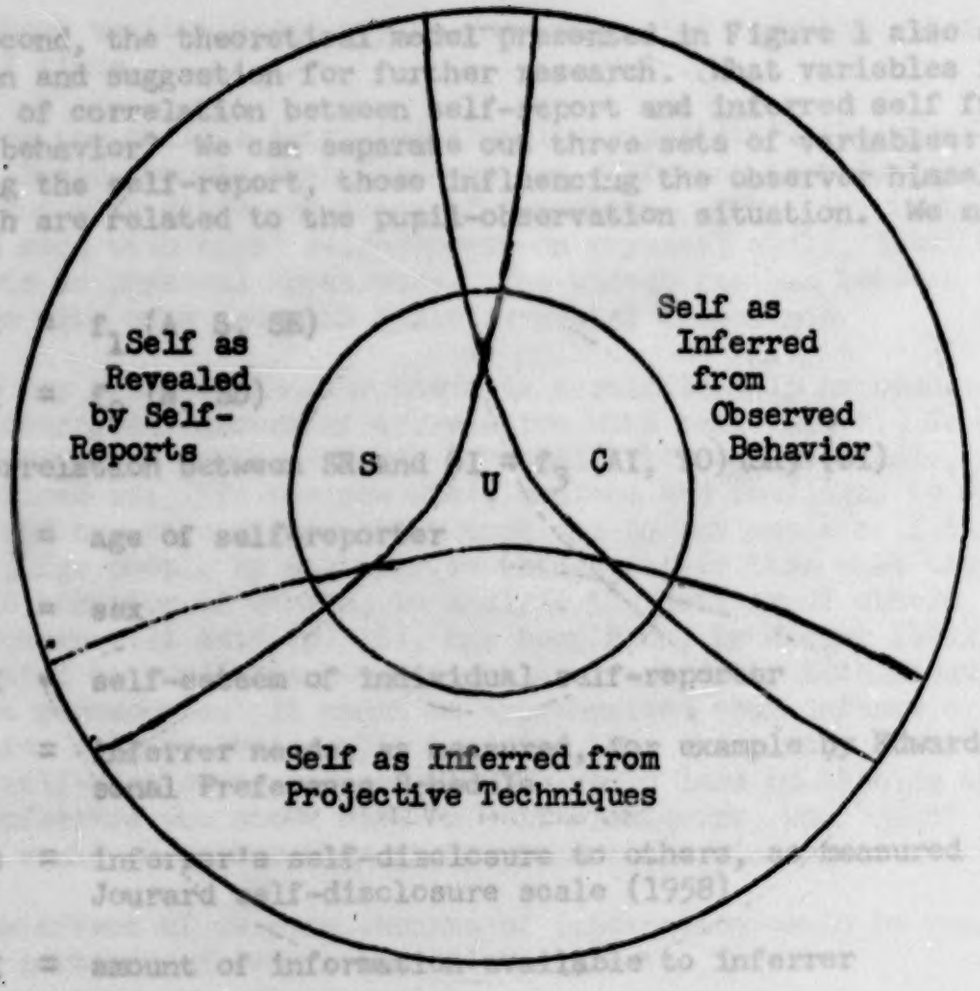

Fig. 1. A tentative model depicting three facets of the self.

Key: S-C is the Self Concept, that highly organized, central core of the self.

U is the part of the self unavailable to public study.

Second, the theoretical model presented in Figure 1 also serves as an implication and suggestion for further research. What variables influence the amount of correlation between self-report and inferred self from observation of behavior? We can separate out three sets of variables: those influencing the self-report, those influencing the observer himself, and those which are related to the pupil-observation situation. We might say that:

$SR = f_1 (A, S, SE)$

$OI = f_2 (N, SD)$

Correlation between SR and OI = f_3 (AI, TO)(SR) (OI)

where A = age of self-reporter

S = sex

SE = self-esteem of individual self-reporter

N = inferrer needs, as measured, for example by Edwards Personal Preference Schedule

SD = inferrer's self-disclosure to others, as measured by Jourard self-disclosure scale (1958)

AI = amount of information available to inferrer

TO = type of observation, that is, participant or non-participant, ecological or structured

Of course, the above does not represent a complete list of variables, but does suggest an organization for further research. One can address oneself to such questions as: Does the self-report of pupils vary with age and sex in any predictable pattern? Both Perkins' (1958) and Sears' (1960) results suggest they do. Sears reports that these areas of competence are intercorrelated in middle childhood (1960). An unpublished report by Gordon and Spears shows a definite pattern by age and sex on the peer items of the "How I See Myself" scale. Since this scale includes school, peer, physical and "general self" items, it may be that pupils become more variable from item to item at certain ages and less variable at other times. A pilot study of intra-individual variation on this scale, using subjects from third through twelfth grade, is under way.

The sex and age factor may be related in still another way. Do these variables influence youngsters' reports of themselves? What is the relationship between them and the type of behavior which lends itself to observer inference? It may be that one should not expect a high relationship between the self-report and observed behavior of a late adolescent boy, although such an expectation may be held for a pre-adolescent girl. Various areas of the self may be more important to the child at different times as he grows;

thoughts about various areas may be more protected from the outside world at critical times.

Does the child's intra-individual variability influence observer accuracy? Sears reports that teachers had different correlations with boys and girls and with different areas of competence. That is, teacher ratings correlated most with boys' self-reports on physical skill, least with girls' self-reports on physical appearance. The widest gap was between teacher correlation with boys and with girls on mental competence.

We may also ask whether there is a relationship between inferrer need, self-disclosure, and amount of correlation with self-report. Need for intraception as measured by the Edwards Personal Preference Schedule, in which this need is defined as: "To analyze one's motives and feelings, to observe others, to understand how others feel about problems to put one's self in another's place, to judge people by why they do things rather than what they do, to analyze the behavior of others, to analyze the motives of others, to predict how others will act" (p. 11), has been found by Medley (1961) to be highly related to pupil-teacher rapport as measured by both behavioral and self-report procedures. It might be hypothesized that inferrers with a high Nintraception will be better able to infer than those with low scores. Jourard's self-disclosure approach (1958) could lead to testing the hypothesis that inferrers who score high on self-disclosure, who "open" themselves to others, would do better than those who are low disclosers.

The effect of varying amounts of information could be measured, using controlled portions of Debbie and Alice.

The model offers another avenue of research. Training students to infer from behavior and then measuring them against self-report may not be an adequate approach to measurement of learning to infer. If the "access route" idea portrayed in the model is used, a better test might be to compare the experimental and control groups with inferences made by experts. Such an hypothesis might be stated in terms of the gain in correlation of the experimental group when compared with expert clinicians.

Abstract

Three groups of subjects, one experimental and two controls, were given pre- and post-tests on three instruments, Debbie, Alice, and John, to measure their ability to infer the self-reports of these pupils. They were provided with observation data on the three pupils and completed a "How I See Myself" rating scale the way in which they thought the pupils completed the scales. Correlations were computed for each subject with each pupil, and transformed in Fisher's z's. A pseudo-four way analysis of variance design was used to test the effect of training upon ability to infer.

The hypothesis, that training in perceptual theory, observation and analysis of behavior would increase the ability to infer children's self-reports was not substantiated. The study, however, led to the development of two instruments, Debbie and Alice, which reliably detect the differences between individuals in their ability to infer self-reports from mimeographed observational data.

Several theoretical questions concerning the relationships between inferred self-concept from behavior, self-reports and projective techniques, the internal organization of the self and changes of organization with age; along with suggested research to test phases of these theoretical implications, have been discussed. In addition, a shift to the study of the perceiver (inferrer) of behavior as a possible follow-up of this study has been indicated.

Two general leads thus present themselves: further study of the self of the growing child and further study of aspects of the self of the observer which influence his ability to infer the child's self.

It may be that answers from this further research will offer clues to the problem of teacher selection as well as the development of teaching procedures to increase accuracy in inference of the child's self.

Appendix A. Guide for Coding Behavior Records
(Review: Gordon, Teacher as a Guidance Worker, 53-58, 61-69, 108-112, 116-119, 135-162, 187-208)

I. School Expectations

A. What the school expects him to learn.
1. Knowledge (facts, concepts, generalization)
2. Skills (reading, writing, arithmetic, p. e., art, music)
3. Values
(a) Self-understanding
(b) Living with others
(c) Self-direction
(d) Express feeling
(e) Critical thinking
(f) Cleanliness, neatness, order
(g) Moral and Spiritual values
(h) Democracy
(i) Rational authority of the teacher

B. How these are taught
1. classroom climate
2. materials and facilities
3. personnel
4. teaching and evaluating methods
5. school regulations and routines

C. Teacher's behavior and perceptions of subject
1. individual instruction, allowance for individuality
2. approval, recognition, acknowledgement (or reverse)

II. The Peer Society

A. Structure
1. Sex of membership in peer groups
2. Requirements for membership
3. Status arrangement
4. Roles played

B. Activities, language, rituals, clothing

C. Behavior toward adults

D. Behavior toward subject. How do they perceive him?

III. Family Expectations

IV. Biological forces shaping his behavior

A. Level of development (height, weight, secondary sex characteristics, body proportions)

B. Coordination

C. Rate of maturity

D. Health

E. Energy production

V. Self - Processes - Perception

A. Behavior related to perception of self in relation to school work.
   1. defensive behaviors
   2. seeking experiences - participation
   3. actual production of work
   4. use of materials
   5. following rules and regulations
   6. incorporation of values
   7. statements about ability, etc.

B. Behavior related to perception of self in relation to teachers
   1. reactions to approval, recognition, etc.
   2. reactions to disapproval, etc.
   3. initiating contacts with teacher
   4. avoiding contacts with teacher
   5. statements about self and teachers

C. Behavior related to perception of self in relation to peers.
   1. reactions to peer behavior
   2. participation in activities, etc.
   3. initiating contacts with peers
   4. roles played in peer group
   5. statements about self and peers

D. Behavior related to perception of self in relation to his body.
   1. grooming and attire
   2. gross muscular activities
   3. choice of physical activities
   4. performance in physical activities
   5. statements about his body

E. Behavior related to perception of self in relation to his family.

Appendix B. Preliminary Analysis

I. A. What are the academic expectations held by the school for your observee?
   B. What are the behavioral expectations?
   C. How are these conveyed to him?
   D. How is he perceived by his teachers?

II. What are the characteristics of his peer society?

   A. Structure and organization
   B. Activities, rituals, patterns of dress and language
   C. Roles and statuses accorded these roles
   D. Deduce, from A, B, C, the values of his peer society
   E. Describe his immediate peer groups(s) using the above variables

III. Describe the youngster physically:

   A. What is his level of maturity, and what has been his rate of growth?
   B. Has he sufficient energy for school tasks?
   C. How well-coordinated is he?
   D. How does he compare with classmates in height, weight, general body build, etc?
   E. Are there any special physical factors or medical history which should be noted?

   For I, II and III, give support for each generalization by citing references to observations.

IV. Inferences about self-concept

   A. List (without support) those inferences your group feels are most likely to be valid.
   B. List (without support) those you feel <u>might</u> be so, but in which you feel more needs to be known.
   C. List (without support) those about which you have disagreements.
   D. Present a plan for the next series of observations to:

      1. Test inferences listed under A.
      2. Gather data to test B list.
      3. Resolve disagreements on C list.

Appendix C. Final Summary

I. State your group conclusion concerning:

A. The pupil's concepts of himself in relation to school.
B. The pupil's concepts of school work itself.
C. The pupil's concepts of himself in relation to teacher(s).
D. The pupil's concepts of teacher(s).
E. The teacher's(s) appraisal of the pupil's ability, behavior, etc.
F. The pupil's concept of himself in relation to his peers.
G. The pupil's concept of his peers.
H. The peer's concept of the pupil.
I. The peer society (see II of guide for sub-heads).
J. The pupil's concepts of his body.
K. Your conclusions about his body.

II. What does your group see as his over-all self-concept? Does he feel secure? Adequate?

III. How would your group rank in importance to him (hierarchy of values) the self-concepts listed in I? In other words, which are most central, which most peripheral and subject to change?

IV. What does your group see as the important discrepancies between his perceptions and the perceptions of others about him?

. You need not give page reference support for I, but give logical justification for your conclusions in II, III, and IV.

V. As an individual, what would you do as his teacher to utilize the above knowledge in helping him to enhance his self in positive ways in school? (Attach all the individual reports to the group report).

APPENDIX D

Elementary Form

HOW I SEE MYSELF

1.	Nothing gets me too mad	1	2	3	4	5	I get mad easily and explode
2.	I don't stay with things and finish them	1	2	3	4	5	I stay with something till I finish
3.	I'm very good at drawing	1	2	3	4	5	I'm not much good in drawing
4.	I don't like to work on committees, projects	1	2	3	4	5	I like to work with others
5.	I wish I were smaller (taller)	1	2	3	4	5	I'm just the right height
6.	I worry a lot	1	2	3	4	5	I don't worry much
7.	I wish I could do something with my hair	1	2	3	4	5	My hair is nice-looking
8.	Teachers like me	1	2	3	4	5	Teachers don't like me
9.	I've lots of energy	1	2	3	4	5	I haven't much energy
10.	I don't play games very well	1	2	3	4	5	I play games very well
11.	I'm just the right weight	1	2	3	4	5	I wish I were heavier, lighter
12.	The girls don't like me, leave me out	1	2	3	4	5	The girls like me a lot, choose me
13.	I'm very good at speaking before a group	1	2	3	4	5	I'm not much good at speaking before a group
14.	My face is pretty (good looking)	1	2	3	4	5	I wish I were prettier (good looking)
15.	I'm very good in music	1	2	3	4	5	I'm not much good in music
16.	I get along well with teachers	1	2	3	4	5	I don't get along with teachers
17.	I don't like teachers	1	2	3	4	5	I like teachers very much
18.	I don't feel at ease, comfortable inside	1	2	3	4	5	I feel very at east, comfortable inside
19.	I don't like to try new things	1	2	3	4	5	I like to try new things

HOW I SEE MYSELF

20.	I have trouble controlling my feelings	1	2	3	4	5	I can handle my feelings
21.	I do well in school work	1	2	3	4	5	I don't do well in school
22.	I want the boys to like me	1	2	3	4	5	I don't want the boys to like me
23.	I don't like the way I look	1	2	3	4	5	I like the way I look
24.	I don't want the girls to like me	1	2	3	4	5	I want the girls to like me
25.	I'm very healthy	1	2	3	4	5	I get sick a lot
26.	I don't dance well	1	2	3	4	5	I'm a very good dancer
27.	I write well	1	2	3	4	5	I don't write well
28.	I like to work alone	1	2	3	4	5	I don't like to work alone
29.	I use my time well	1	2	3	4	5	I don't know how to plan my time
30.	I'm not much good at making things with my hands	1	2	3	4	5	I'm very good at making things with my hands
31.	I wish I could do something about my skin	1	2	3	4	5	My skin is nice-looking
32.	School isn't interesting to me	1	2	3	4	5	School is very interesting
33.	I don't do arithmetic well	1	2	3	4	5	I'm real good in arithmetic
34.	I'm not as smart as the others	1	2	3	4	5	I'm smarter than most of the others
35.	The boys like me a lot, choose me	1	2	3	4	5	The boys don't like me, leave me out
36.	My clothes are not as I'd like	1	2	3	4	5	My clothes are nice
37.	I like school	1	2	3	4	5	I don't like school
38.	I wish I were built like the others	1	2	3	4	5	I'm happy with the way I am
39.	I don't read well	1	2	3	4	5	I read very well
40.	I don't learn new things easily	1	2	3	4	5	I learn new things easily

REFERENCES

Allport, G. W. "The Trend in Motivational Theory," The American Journal of Orthopsychiatry, 1953, 23, 107-119.

Amatora, Sister Mary. "Validity in Self Evaluation," Educational and Psychological Measurement, 1956, 16, 119-126.

Arbuckle, D. Guidance and Counseling in the Classroom, Boston: Allyn & Bacon, 1957.

Baker, B. O. and Block, J. "Accuracy of Interpersonal Prediction as a Function of Judge and Object Characteristics," Journal of Abnormal and Social Psychology, 1957, 54, 37-43.

Barker, R. and Wright, H. Midwest and Its Children, Evanston, Illinois: Row, Peterson & Co., 1954.

Brandt, R. M. "The Accuracy of Self Estimate: A Measure of Self-Concept Reality," Genetic Psychology Monographs, 1958, 58, 55-99.

Combs, A. W. and Snygg, D. Individual Behavior, (Rev. Ed.), New York: Harper & Bros., 1959.

Combs, A. W. and Soper, D. W. "The Self, its Derivate Terms, and Research," Journal of Individual Psychology, 1957, 13, 134-145.

Combs, A. W., Soper, D. and Courson, C. "The Relationship of Self Concept to Self Report," 1961, unpublished report.

Coopersmith, S. "A Method for Determining Types of Self-Esteem," Journal of Abnormal and Social Psychology, 1959, 59, 87-94.

Cronbach, L. J. "Processes Affecting Scores on 'Understanding of Others' and 'Assumed Similarity,'" Psychological Bulletin, 1955, 52, 177-193.

Crow, W. J. "The Effect of Training Upon Accuracy and Variability in Interpersonal Perception," Journal of Abnormal and Social Psychology, 1957, 55, 355-359.

Davidson, H. H. and Lang, G. "Children's Perceptions of their Teachers' Feelings Toward Them Related to Self-Perception, School Achievement, and Behavior," Journal of Experimental Education, 1960, 29, 107-118.

Gordon, I. J. "Inferring Children's Concepts of Self: Inter-Observer Reliability," 1959, paper read at Southeastern Psychological Association Convention.

Gordon, I. J. "Observing From a Perceptual Viewpoint," Journal of Teacher Education, 1959, 10, 280-284.

Gordon, I. J. The Teacher as a Guidance Worker, New York: Harper & Bros., 1956.

Gordon, I. J. and Combs, A. W. "The Learner: Self and Perception," Review of Educational Research, 1958, 28, 433-444.

Gordon, I. J. and Spears, W. D. "Developmental Patterns of Peer Relations as Revealed by Self-Estimate Data," 1961, paper read at Florida Psychological Association Convention.

Havighurst, R. Lecture on "The Gifted Child," University of Florida Summer Lecture Series, 1959.

Jersild, A. T. In Search of Self, New York: Teachers College, Columbia University, 1952.

Jourard, S. and Lasakow, P. "Some Factors in Self-Disclosure," Journal of Abnormal and Social Psychology, 1958, 56, 91-98.

Kelley, E. Education for What is Real, New York: Harper & Bros., 1947.

Lipsitt, L. P. "A Self-Concept Scale for Children and its Relationship to the Children's Form of the Manifest Anxiety Scale," Child Development, 1958, 29, 463-472.

McNemar, Q. Psychological Statistics, New York: Wiley & Co., 1955.

Medley, D. "Teacher Personality and Teacher-Pupil Rapport," Journal of Teacher Education, 1961, 12, 152-156.

Medley, D. and Mitzell, H. "Some Behavioral Correlates of Teacher Effectiveness," Journal of Educational Psychology, 1959, 50, 239-246.

Moustakas, C. The Teacher and The Child, New York: McGraw-Hill, 1956.

Perkins, H. V., Jr. "Factors Influencing Change in Children's Self-Concepts," Child Development, 1958, 29, 221-230.

Prescott, D. A. The Child in the Educative Process, New York: McGraw-Hill, 1957.

Sears, P. "The Pursuit of Self-Esteem: The Middle-Childhood Years," Newsletter, Division of Developmental Psychology, American Psychological Association, Fall, 1960, 5 pp.

Spears, W. D. "A Variance Design Utilizing a Test for Individual Uniqueness," Unpublished Ph.D. dissertation, George Peabody College for Teachers, 1954.

Strang, R. M. The Adolescent Views Himself, New York: McGraw-Hill, 1957.

Tagiuri, R. and Petrullo, L. (Eds.). Person Perception and Interpersonal Behavior, Stanford, California: Stanford University Press, 1958.